ABSTRACT

Described herein are methods to enhance protein secretion in a host cell. In preferred embodiment, the host cell is a gram-positive microorganism such as a *Bacillus*. In another preferred embodiment, the host cell is a gram-negative microorganism. Preferably the gram-negative microorganism is an *Escherichia coli* or a member of the genus *Pantoaea*. Protein secretion may be enhanced by the overexpression of protein components of the Tat pathway. Alternatively, secretion of foreign proteins can be selectively enhanced by forming a chimeric polypeptide comprising a tat signal sequence and the protein of interest. In a preferred embodiment, the tat signal sequence is selected from phoD or LipA.

GC634-2

Table I. Predicted Twin-Arginine Signal Peptides of *B. subtilis**

protein	signal peptide
AlbB	MSPAQRRILLYILSFIFVIGAVVYFVKSDYLFTLIFIAIAILF
AmyX™	MVSIRRSFEAYVDDMNIITVLIPAEQKEIM
AppB™	MAAYIIRRTLMSIPILLGITLLSFVIMKAAPG
LipA	MKFVKRRIIALVTILMLSVTSLFALQPSAKAAEH
OppB™	MLKYIGRRLVYMIITLFVIVTVTFFLMQAAPG
PbpX	MTSPTRRRTAKRRRRKLNKRGKLLFGLLAVMVCITIWNALHR
PhoD	MAYDSRFDEWVQKLKEESFQNNTFDRRKFIQGAGKIAGLSLGLTIAQSVGAFEV
QcrA	MGGKHDISRRQFLNYTLTGVGGFMAASMLMPMVRFA
SpoIIIJ	MLLKRRIGLLLSMVGVFMLLAGCSSV
TlpA™	MKKTLTTIRRSSIARRLIISFLLILIVPITALSVSAYQS
WapA	MKKRKRRNFKRFIAAFLVLALMISLVPADVLAKST
WprA	MKRRKFSSVVAAVLIFALIFSLFSPGTKAAAAGA
YceA™	MEMFDLEFMRRAFLAGGMIAVMAPILGVYLVLRRQ
YdeJ	MKKRRKICYCNTALLLMILLAGCTDS
YdhF	MRRILSILVFAIMLAGCSSN
YdhK	MSAGKSYRKKMKQRRMNMKISKYALGILMFSLVFVLSACGNNN
YesM™	MKKRVAGWYRRMKIKDKLFVFLSLIMAVSFLFVYSGVQYAFHV
YesW	MRRSCLMIRRRKRMFIAVTLLVLLVMGTSVCPVKAEGA
YfkN™	MRIQKRRTHVENILRILLPPIMILSLILPTPPIHAEES
YkpC	MLRDLGRRVAIAAILSGIILGGMSISLANMP
YkuE	MKKMSRRQFLKGMFGALAAGALTAGGGYGYARYL
YmaC	MRRFLLNVILVLAIVLFLRYVHYSLEPE
YmzC	MFESEAELRRIRIALVWIAVFLLFGACGN
YubF™	MQKYRRRNTVAFTVLAYFTFFAGVFLFSIGLYNADNL
YuiC™	MMLNMIRRLLMTCLFLLAFGTTFLSVSGIEAKDL
YvhJ	MAERVRVRVRKKKKSKRRKILKRIMLLFALALLVVVGLGGYKLY
YwbN	MSDEQKKPEQIHRRDILKWGAMAGAAVAIGASGLGGLAPLVQTAAKP

* Putative twin-arginine signal peptides were identified in two ways. First, the presence of the consensus sequence R-R-X-ϕ-ϕ (ϕ is a hydrophobic residue), immediately in front of an amino-terminal hydrophobic region as predicted with the TopPred2 algorithm (34, 35), was determined. To this purpose, the first 60 residues of all annotated proteins of *B. subtilis* in the SubtiList database (http://bioweb.pasteur.fr/Genolist/Subtilist.html) were used. Second, within the group of twin-arginine membrane sorting signals, cleavable signal peptides were identified with the SignalP algorithm (61, 62). Conserved residues of the twin-arginine consensus sequence (R-R-X-ϕ-ϕ) are indicated in bold. In addition, positively charged residues that could function as a so-called Sec-avoidance signal (54) are indicated in bold and italics. The hydrophobic H-domain is indicated in gray shading. In signal peptides with a predicted signal peptidase I cleavage site, residues from position –3 to –1 relative to the signal peptidase I cleavage site are underlined. Notably, some of these proteins contain one or more putative transmembrane segments elsewhere in the protein (indicated with "TM"), or are putative lipoproteins. Residues forming a so-called lipobox for signal peptidase II cleavage are enlarged in size.

Table II. Plasmids and Strains

Plasmids	Relevant properties	Reference
pUC21	cloning vector; 3.2 kb; Ap^r	63
pJCd1	pUC21 derivative; carrying the *tatCd* gene; 5.4 kb; Ap^r	This work
pJCd2	pUC21 derivative for the disruption of *tatCd*; 6.3 kb; Ap^r; Km^r	This work
pJCy1	pUC21 derivative; carrying the *tatCy* gene; 5.3 kb; Ap^r	This work
pJCy2	pUC21 derivative for the disruption of *tatCy*; 6.5kb; Ap^r; Sp^r	This work
pMutin2	pBR322-based integration vector for *B. subtilis*; containing a multiple cloning site downstream of the P*spac* promoter, and a promoter-less *lacZ*-gene preceded by the RBS of the *spoVG* gene; 8.6 kb; Ap^r; Em^r	31
pMICd1	pMutin2 derivative; carrying the 5' part of the *B. subtilis tatCd* gene	This work
pMICy1	pMutin2 derivative; carrying the 5' part of the *B. subtilis tatCy* gene	This work
pDG792	contains a Km resistance cassette; 4.0 kb; Ap^r, Km^r	64
pDG1726	contains a Sp resistance cassette; 3.9 kb; Ap^r, Sp^r	64
Strains		
E. coli		
MC1061	F^-; *araD*139; Δ (*ara-leu*)7696; Δ (*lac*)X74; *galU*; *galK*; *hsdR*2; *mcrA*; *mcrB*1; *rspL*	65
B. subtilis		
168	*trpC2*	2
Δ*tatCd*	*trpC2*; *tatCd*; Km^r	This work
Δ*tatCy*	*trpC2*; *tatCy*; Sp^r	This work
I*tatCd*	*trpC2*; P*spac-tatCd*; *tatCd-lacZ*; Em^r	This work
I*tatCy*	*trpC2*; P*spac-tatCy*; *tatCy-lacZ*; Em^r	This work
Δ*tatCd*-Δ*tatCy*	*trpC2*; *tatCd*; Km^r; *tatCy*; Sp^r	This work
I*tatCd*-Δ*tatCy*	*trpC2*; P*spac-tatCd*; *tatCd-lacZ*; Em^r; *tatCy*; Sp^r	This work

GC634-2

Table III. β-galactosidase activity (U/OD_{600})*.

strain	LPDM	MM	SSM	TY
168	0	0.1 ± 0.1	0.3 ± 0.2	0.6 ± 0.2
ItatCd	1.1 ± 0.7	0.1 ± 0.1	0.3 ± 0.2	0.5 ± 0.2
ItatCy	6.1 ± 2.5	10.0 ± 3.6	4.0 ± 2.0	13.2 ± 5.5

* To investigate the transcription of the *tatCd* and *tatCy* genes, cells of *B. subtilis* I*tatCd* (*tatCd-lacZ*), I*tatCy* (*tatCy-lacZ*) or the parental strain 168 (no *lacZ* gene fusion) were grown for 10 hours in LPDM, MM, SSM or TY medium after dilution from an overnight culture. Samples for β-galactosidase activity determinations were taken at hourly intervals, starting 4 hours after dilution from the overnight culture. As the β-galactosidase activities showed little variation during the entire period of sampling, average values were determined. The numbers in the table represent average values from 3 different experiments. Note that HPDM medium was used for the overnight culture of cells grown in LPDM medium, while overnight cultures of cells grown in MM, SSM or TY medium were prepared with the respective media.

GC634-2

Table IV. Twin-Arginine Signal Peptides of PhoD and PhoD-like proteins*

protein	signal peptide
PhoD (Bsu)	MAYDSRFDEWVQKLKEESFQNNTFDRRK**FI**QGAGKIAGLSLGLTIAQSVGAFEV
SP1 (Sco)	MTPANHQAPTSAPSPAPSQSSHAPELRAAARSLGRRR**FL**TVTGAAAALAFAVNLPAAGTASAAEL
SP2 (Sco)	MAPTGRPSALAEHAFSPHDAVLGAAARHLGRRR**FL**TVTAAAAALAFSTNLPARGAVAAPE
SP3 (Sco)	MTSRHRASENSRTPSRRTVVKAAAAGAVLAVAPLAAVALPAGAADAAPA
SP4 (Ste)	MTPAARPSQHAPELRAAARHLGRRR**FL**TVTGAAAALAFAVNLPAVGTAAAAEL

* Homologues of *B. subtilis* PhoD were identified by amino acid sequence similarity searches in GenBank using the BLAST algorithm. SP1 (Sco), gene SCC75A.32c of *Streptomyces coelicolor* (CAB61732); SP2 (Sco), gene SCF43A.18 of *S. coelicolor* (CAB48905); SP3 (Sco), gene SC4G6.37 of *S. coelicolor* (CAB51460), and SP4, *phoD* gene of *Streptomyces tendae* (CAB62565). GenBank accession numbers are indicated in parentheses. Conserved residues of the twin-arginine consensus sequence are indicated in bold. The hydrophobic H-region is indicated in in gray shading. Signal peptidase I recognition sequences predicted with the SignalP algorithm (61, 62) are underlined.

TABLE 5- *Plasmids and Strains*

Plasmids	Relevant properties	Reference
pAR3	pACYC184 derived plasmid carrying the *araB* promoter operator and the *araC* repressor gene from *Salmonella typhimurium*; Cm^r [a]	25
pAR3*phoD*	pAR3 derivative; carrying the *phoD* gene; Cm^r	This work
pAR3*phoD-lacZ*	pAR3 derivative; carrying a fusion gene consisting of the signal sequence region of *phoD* and *lacZ*; Cm^r	This work
pQE9	pBR322-based vector for IPTG-inducible synthesis of His_6-tagged proteins; Ap^r	Qiagen
pREP4	plasmid; containing *lacI*^{q} repressor gene; Km^r	Qiagen
pORI24	plasmid; replicates only in *E. coli rep*$^+$ strains; Tc^r	37
pMUTIN2	pBR322-based integration vector for *B. subtilis*; containing a multiple cloning site downstream of the P*spac* promoter, and a promoter-less *lacZ*-gene preceded by the RBS of the *spoVG* gene; Ap^r; Em^r	38
pMUTIN2*bla-phoD*	PMUTIN2 derivative; carrying a fusion gene consisting of signal sequence region of *bla* and *phoD*	This work
pQE9*tatA$_d$/C$_d$*	pQE9 derivative; carrying the *B. subtilis tatA$_d$/C$_d$* genes	This work
pFAT44	pMAK705 (Hamilton *et al.*, 1989) derivative plasmid containing in frame deletion of *E. coli tatE*	7
pFAT126	pMAK705 derivative plasmid containing in frame deletion of *E. coli tatABCD*	39
Strains		
E. coli		
TG1	F$^-$ *araD*139 Δ(*ara-leu*)7696 Δ (*lac*)X74 *galU galK hsdR2 mcrA mcrB*1 *rspL*	40
TG1 Δ*tatABCE*	TG1 Δ*tatABCE*	This work
B. subtilis		
168	*trpC2*	13

[a] Cm^r, chloramphenicol resistance marker; Ap^r, ampicillin resistance marker; Km^r, kanamycin resistance marker; Tc^r, tetracycline resistance marker; Em^r, erythromycin resistance marker

TABLE 6 - *Localisation of ß-galactosidase activity in E. coli TG1(pAR3phoD-lacZ) strains.*

To investigate the translocation of the hybrid protein consisting of SP_{PhoD} and LacZ, cells of *E. coli* strains were grown in TY medium to exponential growth. Samples for ß-galactosidase activity determinations were taken from supernatants of lysozyme treated cells representing periplasmic activity and spheroplasts representing cell bound activity. Experiments were carried out with duplicated cultures. +/-, standard deviation.

	LacZ activity (units/OD_{600})			
strain	cell bound	periplasmic	total activity	% export
TG1(pAR3*phoD-lacZ*)	1108 +/- 201	67 +/- 5	1175	6.4 +/- 3,4
TG1(pAR3*phoD*-lacZ, pREP4, pQE9*tatA$_d$/C$_d$*)	226 +/- 11	94 +/- 2	320	29.4 +/- 0.4
TG1 Δ*tatABCE* (pAR3*phoD-lacZ*, pREP4, pQE9*tatA$_d$/C$_d$*)	278 +/- 8	39 +/- 5	317	12.5 +/- 0.9